Fourier Deconvolution of the Amide I Raman Band of Proteins as Related to Conformation

HEINO SUSI* and D. MICHAEL BYLER†
Eastern Regional Research Center, U.S. Department of Agriculture, Philadelphia, Pennsylvania 19118

Fourier deconvolution has been employed to enhance the resolution of the amide I Raman band of nine proteins found in milk and/or other foods. The broad band was resolved into several components. The overall shape of the amide I Raman band of proteins was found to be nearly Gaussian or to be composed of Gaussian components. A Gaussian function was therefore used for deconvolution. The results obtained were more detailed than those obtained with the Lorentzian approximation usually employed. The resolved band components were assigned to specific protein conformations. The frequencies and assignments are in good agreement with previous Raman work based on entirely different procedures. The band areas of the resolved components appear to reflect the fraction of any given conformation in a protein. Semiquantitative estimations of protein conformation are in reasonable agreement with data obtained by x-ray diffraction and by infrared methods.

Index Headings: **Raman spectra; Fourier deconvolution; Proteins.**

Received 7 September 1987; revision received 28 December 1987.
* The Editor regrets to announce that Dr. Heino Susi died on 2 October 1987.
† Author to whom correspondence should be sent.

INTRODUCTION

Several studies of protein conformation by Fourier deconvolution of the infrared amide I band have been previously carried out.[1-3] Raman studies of protein confor-

mation, on the other hand, have been primarily concerned with deducing information from unresolved amide I and amide III bands by peak intensity measurements[4,5] or by correlation studies and curve fitting.[6-8] The present study applies deconvolution techniques in conjunction with curve fitting methods to the Raman spectra of nine lyophilized proteins in a manner similar to that previously applied to FT-IR spectra of proteins in D_2O solution.[2-3] The main drawbacks of Raman spectra, as compared with FT-IR spectra in this context, are the lower obtainable signal-to-noise ratio, the lower inherent intensity of the amide I Raman band, and the frequently observed background fluorescence. One of the first reports on deconvolved Raman spectra was on polyvinyl chloride.[9] Deconvolved Raman spectra of some proteins have also been reported,[10,11] but the mathematical approach was different from the one applied in this investigation. No detailed assignments of the resolved band components, and no quantitative or semiquantitative conformation studies, appear to have been carried out in these previous investigations.[10,11]

Detailed assignments for deconvolved infrared components of amide I bands of proteins in D_2O solution were previously reported.[3] Although no strict selection rules apply, Raman assignments for lyophilized proteins require a separate study because of (1) variations induced by the change of state, (2) changes caused by the deuteration of backbone amide groups in the previous infrared studies,[1-3] and (3) differences in relative intensities. Thus, the amide II band is of medium intensity in infrared spectra, but is not easily observed in the Raman effect. Similarly, the amide I components associated with β-strands have different relative intensities in the two kinds of spectra.[1-7]

EXPERIMENTAL

Raman spectra of lyophilized proteins sealed in melting point capillaries were obtained from 1450 to 1750 cm^{-1} on a Spex 1401 spectrometer, with the 514.5-nm line of a Spectra-Physics Model 165-03 argon-ion laser used for excitation. The laser power at the sample was ~250 mW. The spectral slit width was about 4 cm^{-1}. The instrument was equipped with a stepping motor controlled by a Spex Datamate computer. One data point was recorded per 1.0 cm^{-1}. Depending on the noise level observed with a particular protein, 6–10 scans were signal averaged and smoothed by a 9-point Golay-Savitzky procedure.[12] For noisy samples 13-point smoothing was tried; it neither improved nor imperiled the information obtained.

The digitized spectra were deconvolved with the help of a slightly modified program (#LI) of the National Research Council of Canada.[13] As compared with the previous programs written by Kauppinen *et al.*,[14] the new version permits different assumptions regarding the shape of the original bands. It was found experimentally that the overall shape of the amide I bands of the proteins are almost Gaussian, or are best fitted with Gaussian components. Figure 1 shows the amide I region of the Raman spectrum of the lyophilized protein α-lactalbumin as an example. A Gaussian form was therefore assumed for the deconvolution of unresolved components as well, because it is impossible to produce an overall Gaussian shape from Lorentzian components. Conservative values were selected for the deconvolution parameters.[13-15] A halfwidth at half-height (HWHH) of 6.5 cm^{-1} was chosen by analogy with previous infrared work in this laboratory and elsewhere.[1,2] The resolution factor,[14] K, was varied between 2.0 and 2.6. Although Gaussian deconvolution leads in this case to better resolution than does the usual Lorentzian assumption, the value of K becomes very critical. We took care to avoid overdeconvolution, as suggested by Mantsch *et al.*,[15] by recording a flat portion of the spectrum on the high-frequency side of the amide I band and checking this region for undue noise and artifacts. For consistency, deconvolution was always carried out over the same range, 1580–1750 cm^{-1}. Figure 2 shows the amide I region of the protein α-lactalbumin as deconvolved with HWHH = 6.5 cm^{-1} and K values ranging from 2.0 to 2.6. Values of 2.2 to 2.4 were selected for the final studies—again close to the values used in previous infrared studies.[1-3]

The samples of β-casein, α-lactalbumin, and β-lactoglobulin (all from bovine milk) were obtained through the courtesy of Dr. Harold M. Farrell, Jr., and Dr. Marvin P. Thompson of this research center. The remaining samples were from the Sigma Chemical Company: albumin (bovine serum), A-0281; carbonic anhydrase (bovine pancreas), C-7500; immunoglobulin G (bovine), I-5506; lysozyme (hen egg white), L-6876; ribonuclease A (bovine pancreas), R-5500; and ribonuclease S (bovine pancreas), R-6000.

RESULTS AND DISCUSSION

Assignments of Band Components. Figure 3 shows the original Raman spectrum, the deconvolved spectrum, and the applied curve fitting for α-lactalbumin. The deconvolved spectra are easily fitted with Gaussian components, as previously described.[2] Figure 4 shows the deconvolved amide I spectra of four proteins, where assignments of characteristic components can be made in a relatively simple manner by comparison with structural data obtained from x-ray diffraction studies[16,17] and infrared studies.[2] The deconvolved spectra were curve fitted as described,[2,3] in order to permit an estimation of relative integrated intensities as related to conformation (see below). The strongest component of immunoglobulin (Fig. 4A) appears at 1673 cm^{-1} and is undoubtedly associated with the β-structure, because the protein is primarily in this conformation.[2,16] Similarly, the strongest component in the spectrum of albumin (4B) appears at 1657 cm^{-1} and is associated with the predominant helical structure of this molecule.[2] Carbonic anhydrase (4C) has a strong β-component at 1675 and a medium-intensity "unordered" component at 1662 cm^{-1}, while ribonuclease S (4D) shows major components at 1670 (β), 1655 (helix), and 1689 cm^{-1} (turns). These assignments (except for the turns) are in good agreement with the ones by Alix *et al.* that were derived by an entirely different computational procedure, the Raman intensity profile (RIP) approach.[7,8] The other bands are assigned by analogy with infrared spectra of proteins[2] and Raman spectra of amino acids.[18] Thus, the two low-frequency bands (near 1604 and 1614 cm^{-1}) are caused by aromatic

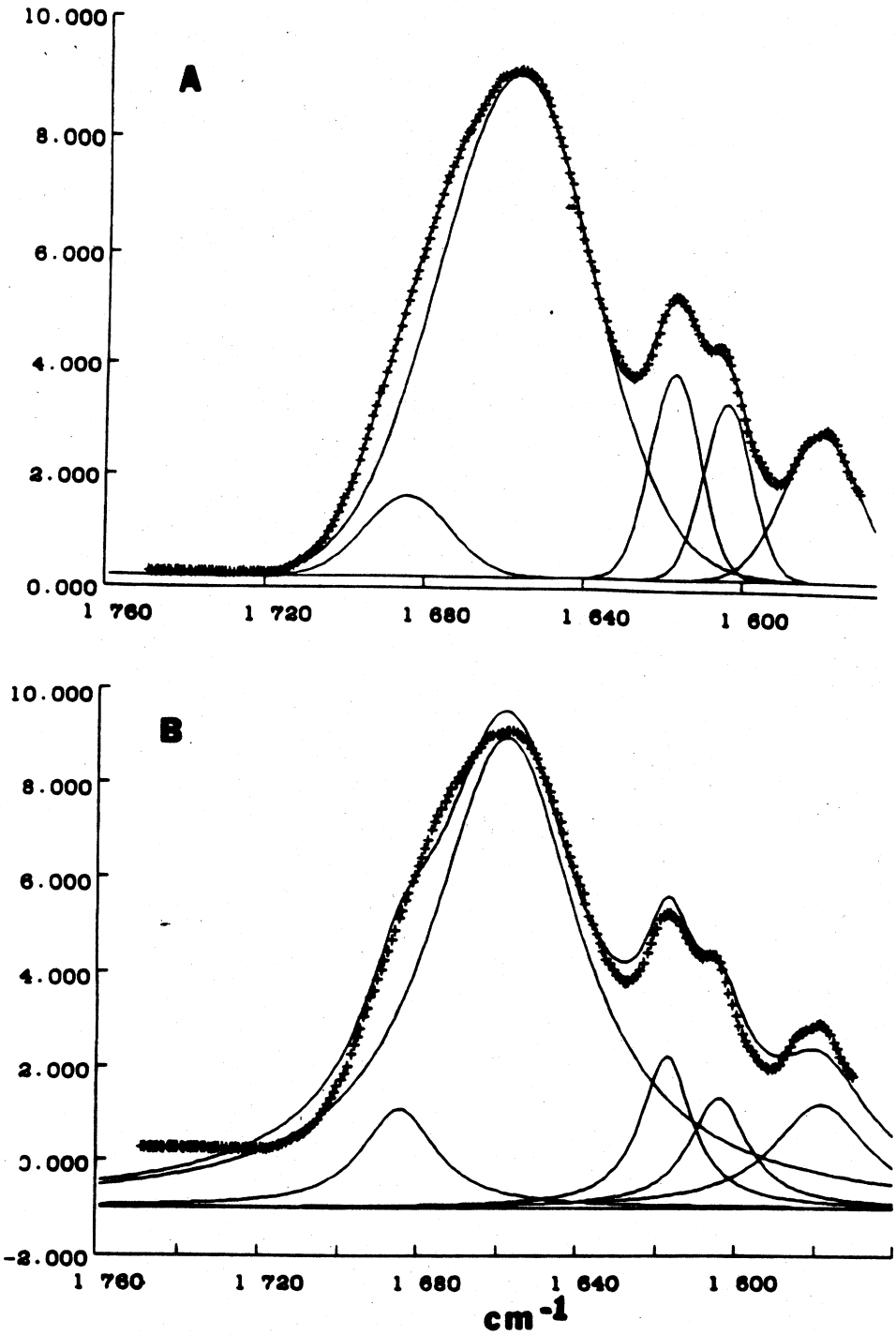

FIG. 1. Original undeconvolved spectrum of α-lactalbumin fitted with Gaussian components (A) and Lorentzian components (B) in the amide I region. In (A) the sum of the components coincides with the experimental curve within experimental error.

side-chains of the individual amino acids, a band close to 1644 cm^{-1} is probably associated with water of hydration bound to the proteins, a weak band close to 1632 cm^{-1} is assigned to the other β-structure band (which is very strong in the infrared but is usually not observed in nondeconvolved Raman spectra[4-8]), and the high-frequency bands are assigned to "turns."[2] The frequencies are, in general, quite close to corresponding infrared frequencies observed in D_2O solution,[2] with the important exception of the band for "unordered" segments, which shifts from ~1645 cm^{-1} in D_2O solution to about 1660 cm^{-1} in the Raman spectra of solids. This decrease in frequency may result in part because some unordered segments of the dissolved protein become surrounded by water molecules, with the result that hydrogen bonds between the peptide C=O and H-O of water replace some of those linking C=O to neighboring H-N of the peptide backbone.[2,3] (In D_2O solution H-O and H-N groups are, of course, replaced by D-O and D-N, respectively.)

Figure 5 shows the spectra of four more proteins in the same spectral region. The assignments are summarized in Table I. It is not possible to resolve all compo-

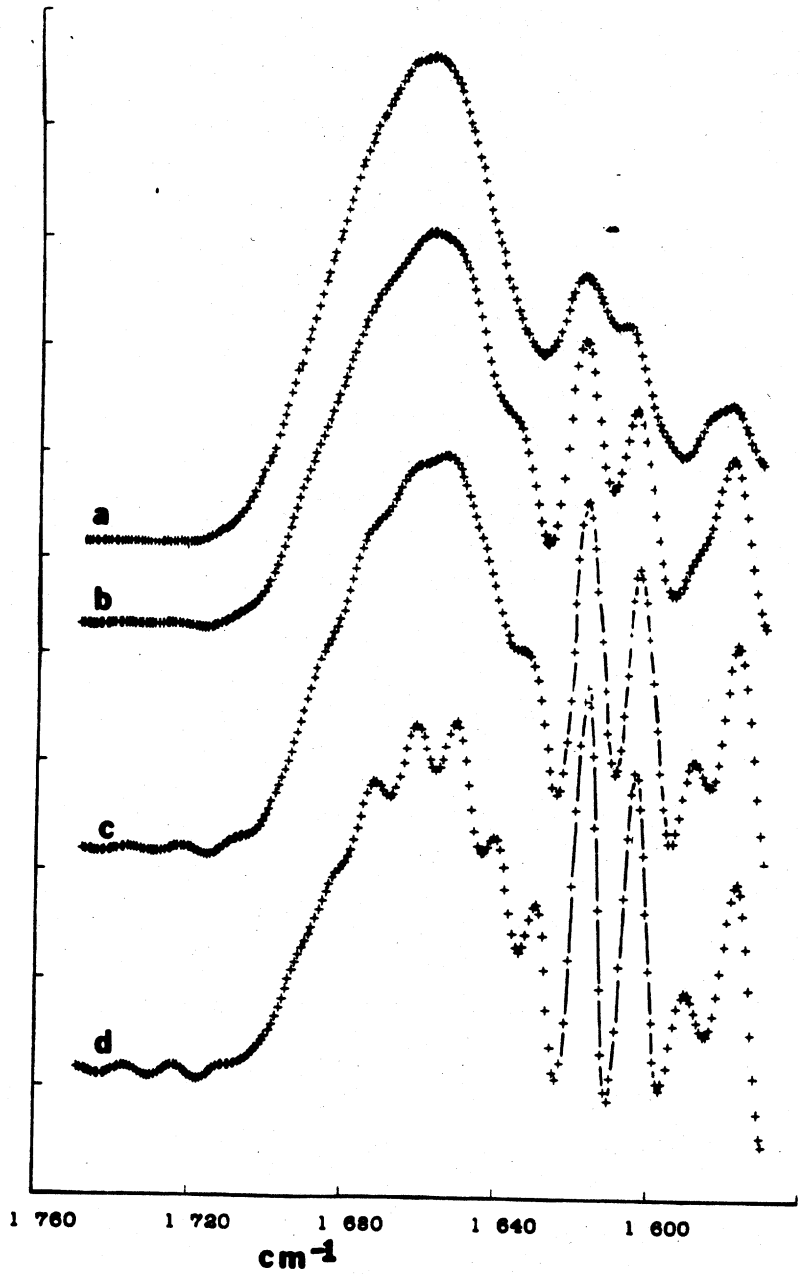

FIG. 2. Deconvolved Raman spectrum, with the use of a Gaussian deconvolution function and different resolution enhancement factors (see text) of α-lactalbumin. Assumed halfwidth at half-height (HWHH): 6.5 cm^{-1}. Resolution factor (K values): (a) 2.0, (b) 2.2, (c) 2.3, (d) 2.5.

nents of all amide I Raman bands by deconvolution, because some band centers are very close to each other and the resolution obtained is not as high as that achieved in previous studies with FT-IR spectra.[1-3] Thus, in β-lactoglobulin (Fig. 5C) the components at 1655 and 1662 cm^{-1} (helical and unordered sections) are not resolved by deconvolution. They are inserted by a comparison with all the other spectra, i.e., by consideration of the collected data given in Table I. We are thus applying a characteristic frequency approach similar to the one previously developed for infrared active amide I band components.[2,3]

Table I summarizes the assignments of all observed band components. The estimated frequency accuracy is about ± 1 cm^{-1}; the range of the observed values from the mean frequency of each characteristic band is somewhat larger (Table I). The positions of components which were not actually resolved by deconvolution are given in parentheses. Table II compares the characteristic mean amide I frequencies obtained in this study with corresponding values obtained by the Raman intensity profile (RIP) method of Alix *et al.*[7,8] and with values obtained by infrared spectroscopy of D_2O solutions.[2] The agreement with the three Raman amide I components reported by Alix and co-workers centering around 1672, 1661 and 1655 cm^{-1} is remarkably good, particularly in view of the fact that the two sets of data were calculated by entirely different mathematical procedures and from completely independent sets of empirical spectra. The major difference between the two sets of Raman assignments concerns the bands for "turns," observed around 1683, 1689, and 1696 cm^{-1} by the Fourier deconvolution

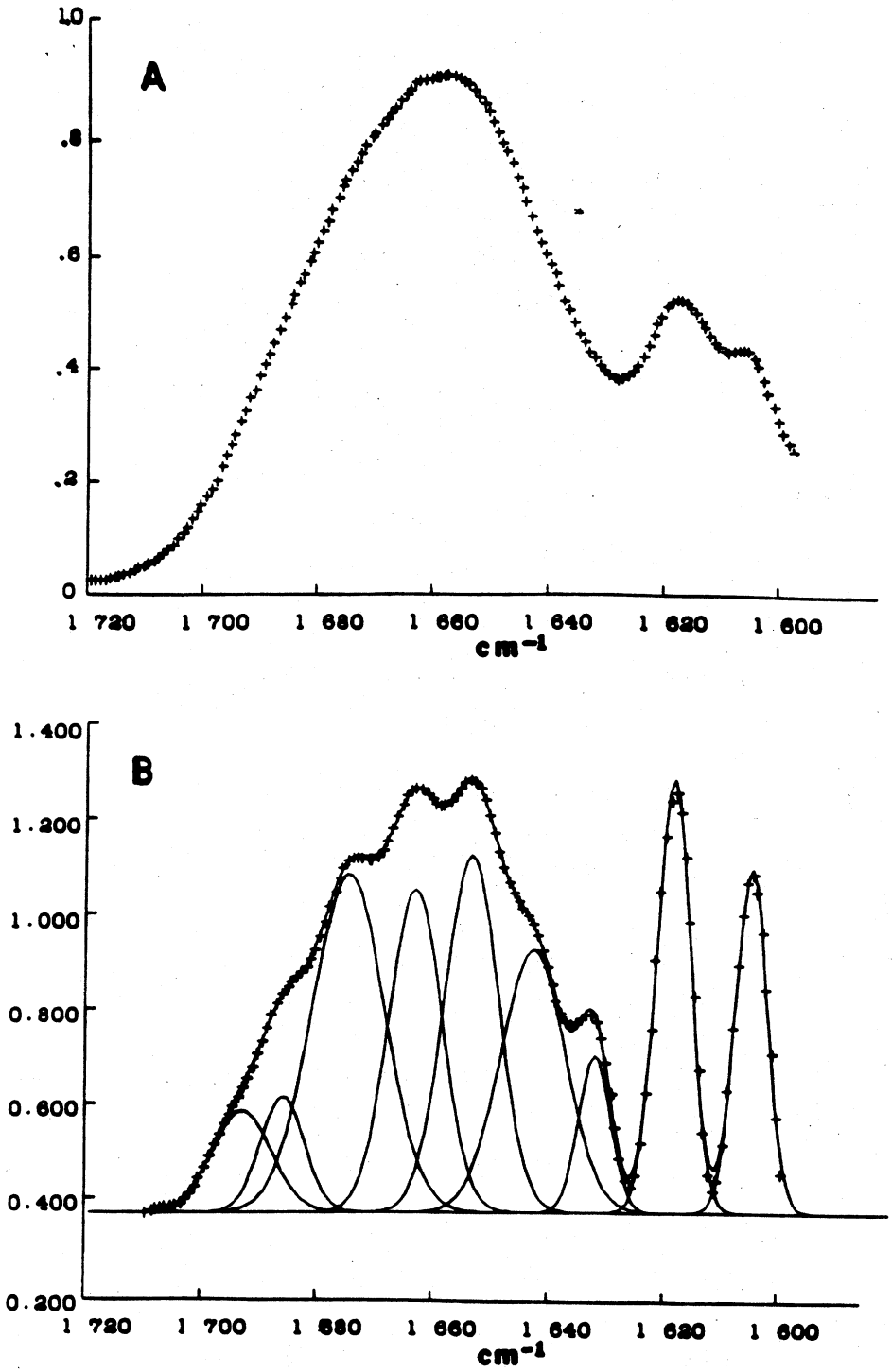

FIG. 3. Original Raman spectrum (above) and deconvolved spectrum fitted with Gaussian components (below) of α-lactalbumin. Deconvolution carried out with a Gaussian function. $K = 2.4$; HWHH = 6.5.

method, but not reported by the RIP method.[7,8] Some of these high-frequency component bands are clearly discernible as shoulders even in the original spectra (see, for example, Fig. 1). Neglect of these components might have caused the irregular shape on the high-frequency side of the reported amide I RIP components.[7] A second difference in the results of these two methods is found on the low-frequency side of the amide I band. Deconvolution resolves a weak component around 1632 cm^{-1} (Table I and Figs. 4–5) which is not observable in the original spectra.[4,5] This weak band can be assigned to the β-structure component, which is strong in infrared but very weak in the Raman effect.[2,4]

The major difference, in comparison to the infrared data, is confined to the band associated with "unordered" segments, because, as discussed above, in solution some peptide groups of these segments are probably hydrogen bonded to the solvent and not to other peptide groups, as in the solid.[2] In D_2O solution these groups absorb at around 1645 cm^{-1} in the infrared.[2] In the Raman spectra

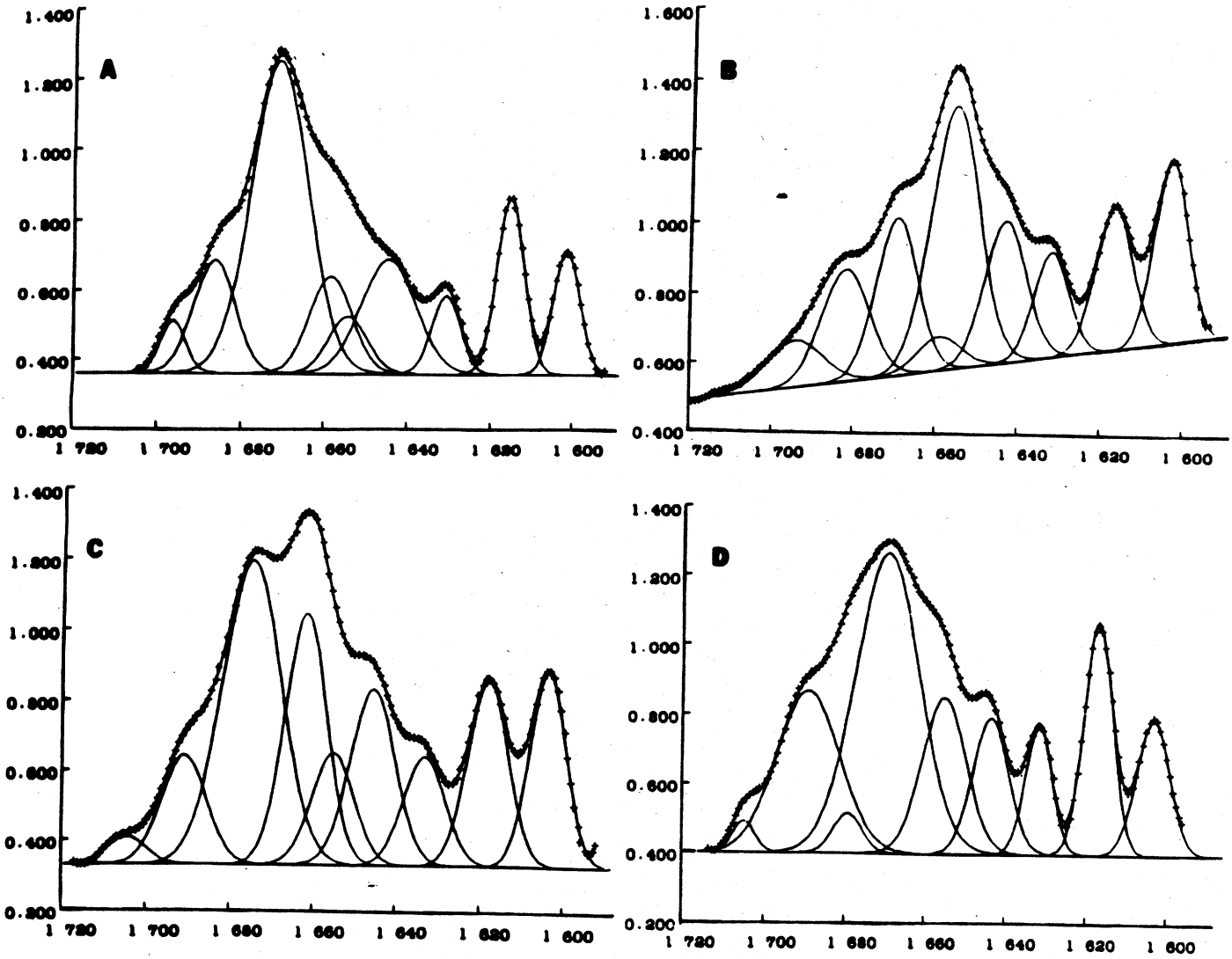

FIG. 4. Deconvolved and curve-fitted amide I Raman bands of four proteins. Gaussian functions used throughout: (A) immunoglobulin G; (B) albumin (serum); (C) carbonic anhydrase; (D) ribonuclease S.

of lyophilized proteins the analogous component appears near 1661 cm^{-1}. Infrared spectra of D_2O solutions of most proteins[2] also have a band around 1663 cm^{-1}, due to turns. The Raman counterpart of this infrared band probably contributes to the overall intensity of the 1661-cm^{-1} Raman component, as indicated in Table II. Thus, the band originally assigned to "undefined" conformations by the RIP method[7,8] is probably in part also associated with turns.

Estimation of Protein Conformation. Although the band

TABLE I. **Raman frequencies (cm^{-1}) of deconvolved amide I bands for lyophilized proteins.**[a]

Protein	Turns			β-Strand	Undef./ turns	α-Helix	[H_2O]	β-Strand
Albumin (serum)	1695	—	1683	1671	(1660)	1657	1644	1633
Carbonic anhydrase	—	1691	—	1675	1662	(1655)	1646	1633
α-Chymotrypsin	—	1690	1681	1668	(1660)	1655	1645	1632
Immunoglobulin G	1698	1688	—	1673	1660	(1656)	1646	1633
Insulin	—	(1688)	1680	(1671)	—	1658	1644	1632
α-Lactalbumin	1693	—	1686	1674	1663	1653	1641	1631
β-Lactoglobulin	1699	1688	—	(1671)	(1662)	(1655)	1645	1632
Lysozyme	1695	—	1682	1671	(1662)	1658	1644	1633
Ribonuclease A	(1696)	—	1685	1671	1661	(1657)	(1646)	1633
Ribonuclease S	—	1689	—	1670	—	1655	1644	1632
Average	1696	1689	1683	1672	1661	1656	1645	1632
Maximum deviation (±)	3	4	3	4	2	3	4	1

[a] Values enclosed in parentheses are for bands which were not resolved by deconvolution but which were obviously necessary for proper curve fitting.

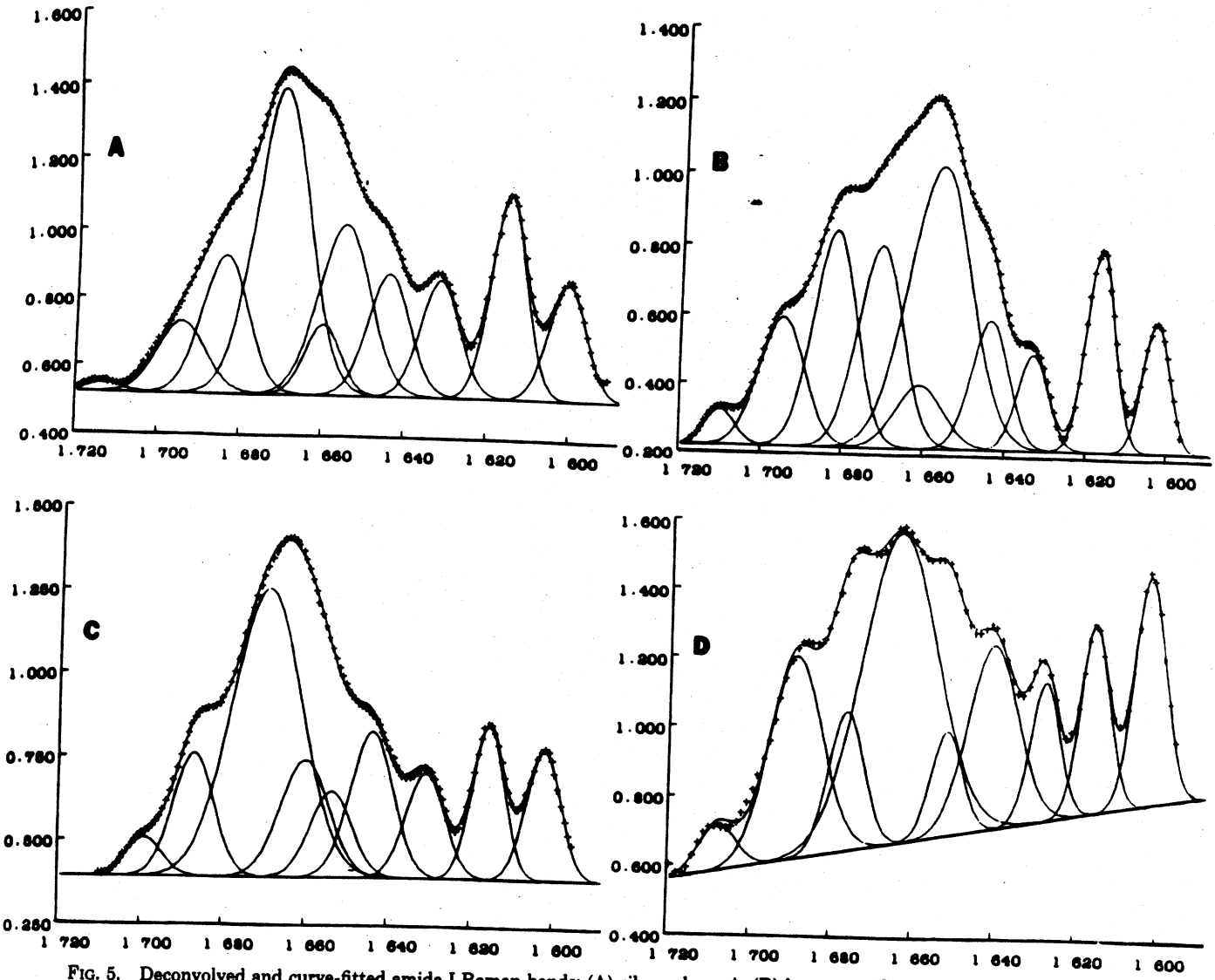

FIG. 5. Deconvolved and curve-fitted amide I Raman bands: (A) ribonuclease A; (B) lysozyme; (C) β-lactoglobulin; (D) β-casein.

fitting procedure represented by the spectra in Figs. 4 and 5 is based in the case of a few bands merely on characteristic frequencies (see previous section), and not on observed resolved components, it is interesting to compare the protein conformation estimated from Raman data (such as % helix, % β-structure, etc.) with corresponding values obtained by infrared spectroscopy[2] and by x-ray diffraction.[16,17] The quantitative evaluations were carried out exactly as previously described for deconvolved infrared spectra.[2] Table III lists some comparative values. The agreement is quite reasonable. This suggests that the scattering intensity of band components associated with various conformational substructures is approximately equal (at least for the amide I bands), just as the integrated intensity of the infrared components appears to remain about the same from one conformation to another.[2] This observation furnishes some hope that, in principle, quite accurate conforma-

TABLE II. Comparison of characteristic amide I frequencies (cm^{-1}) as obtained by different techniques.

Technique	β-Strands		Helix	Undefined and turns		Turns			
FT-IR (FD/CF)[a]	1631[b]	1675	1653	1645	1663	1670[c]	1683	1689	1694
Raman (FD/CF)[d]	1632	1672	1656	1661		—	1683	1689	1696
Raman (RIP)[e]	—	1673	1652	1660		—	—	—	—

[a] FT-IR, Fourier deconvolved and curve fit—average values for 19 proteins in D_2O solution (Refs. 2, 19).
[b] Average value of three characteristic FT-IR β-bands (Refs. 2, 19).
[c] This weak FT-IR component is observed only for four of the 19 proteins surveyed (Refs. 2, 19). For these proteins, the weak Raman counterpart to this FT-IR band will constitute but a minor fraction of the observed area of the high-frequency β-strand component.
[d] Raman, Fourier deconvolved and curve fit—average values for 10 lyophilized, undeuterated proteins (see Table I).
[e] Raman Intensity Profile method (Refs. 7, 8). (Uncertainty, ~±2 cm^{-1}.)

TABLE III. Estimation of protein conformation by different methods.[a]

Protein	% Helix			% β-strand			% Other		
	R	FT-IR	x-ray	R	FT-IR	x-ray	R	FT-IR	x-ray
Albumin (serum)	39	47	—	32	28	—	29	25	—
Carbonic anhydrase	11	13	16	51	49	45	38	38	39
Immunoglobulin G	8	9	3	67	76	67	27	15	30
α-Lactalbumin	31	33	—	36	41	—	33	26	—
β-Lactoglobulin	10	10	7	54	50	51	36	40	42
Lysozyme	43	41	45	25	21	19	32	38	36
Ribonuclease A	21	21	22	50	50	46	29	29	32
Ribonuclease S	17	25	23	56	50	53	27	25	24
β-Casein[b]	(7	7	—)	(19	21	—)	(74	72	—)

[a] Raman, this work; FT-IR, Ref. 2 (except for albumin and β-casein, Ref. 21); x-ray data, Ref. 16 (except for β-lactoglobulin, Ref. 17).
[b] β-Casein has never been crystallized. The data represent approximate amounts of local order in an overall unordered structure (cf. Refs. 20, 21).

tional studies are possible by deconvolved Raman spectra if the difficulties associated with noise and fluorescence can be overcome.

The protein β-casein constitutes a special case. Caseins are milk proteins which no one has been able to crystallize. Careful circular dichroism[20] and more detailed Raman studies,[22] as well as conformational predictions based on amino acid sequences,[20] all suggest that these proteins have relatively little periodic secondary structure. Values for helix-content and β-content of casein must therefore be regarded as somewhat uncertain, in comparison to those for typical globular proteins. Clearly the most intense component is the "aperiodic" band at 1661 cm^{-1}; nevertheless, the Raman results do imply that this protein is not totally "structureless."

We conclude, therefore, that Fourier deconvolution of the amide I Raman band combined with iterative curve fitting yields useful semiquantitative information on the conformation of lyophilized proteins. The results are in good agreement with results from other Raman methods as well as with those from x-ray and infrared studies.

Note added in proof: After this manuscript was accepted for publication, we became aware of some recent papers in which protein secondary structure was estimated from the integrated intensities of amide I Raman bands deconvolved by the constrained, iterative method (see Ref. 10):

1. G. J. Thomas, Jr., Spectrochim. Acta **41A**, 217 (1985).
2. B. Prescott, V. Renugopalakrishnan, M. J. Glimcher, A Bhushan, and G. J. Thomas, Jr., Biochemistry **25**, 2792 (1986).
3. G. J. Thomas, Jr., B. Prescott, J. M. Benevides, and M. A. Weiss, Biochemistry **25**, 6768 (1986).
4. N. L. Incardona, B. Prescott, D. Sargent, O. P. Lambda, and G. J. Thomas, Jr., Biochemistry **26**, 1532 (1987).

ACKNOWLEDGMENTS

The authors are grateful to Dr. R. Norman Jones, Dr. Henry H. Mantsch, and Dr. Douglas J. Moffatt of the National Research Council of Canada for an early version of the computer program #LI, and to Professor Alain J. P. Alix of the Laboratoire de Spectroscopie Biomoleculaire (ISERM U314), Reims Cedex, France, for preprints of some of his publications. We also thank Kevin Newman, Thomas M. Muller, and Stephen McGady for assistance with the experimental measurements.

1. W.-J. Yang, P. R. Griffiths, D. M. Byler, and H. Susi, Appl. Spectrosc. **39**, 282 (1985).
2. D. M. Byler and H. Susi, Biopolymers **25**, 469 (1986).
3. H. Susi and D. M. Byler, Meth. Enzymol. **130**, 290 (1986).
4. J. L. Lippert, D. Tyminski, and P. J. Desmeules, J. Am. Chem. Soc. **92**, 7075 (1976).
5. M. Pézolet, M. Pigeon-Gosselin, and L. Coulombe, Biochim. Biophys. Acta **453**, 502 (1976).
6. R. M. Williams, Meth. Enzymol. **130**, 311 (1986).
7. A. J. P. Alix, M. Berjot, and J. Marx, "Determination of the Secondary Structure of Proteins," in *Spectroscopy of Biological Molecules*, A. J. P. Alix, L. Bernard, and M. Manfait, Eds. (John Wiley & Sons, Chichester, 1985), p. 149.
8. M. Berjot, J. Marx, and A. J. P. Alix, J. Raman Spectrosc. **18**, 289 (1987).
9. J. King and D. I. Bower, *Proceedings of the VIIth International Conference on Raman Spectroscopy*, Ottawa, Canada (Elsevier-North Holland, New York, 1980), p. 242.
10. G. J. Thomas, Jr. and D. A. Agard, Biophys. J. **46**, 763 (1984).
11. F. Ni and H. A. Scheraga, J. Raman Spectrosc. **16**, 337 (1985).
12. M. J. E. Golay and A. Savitzky, Anal. Chem. **36**, 1625 (1964).
13. D. J. Moffatt, J. K. Kauppinen, H. H. Mantsch, and R. N. Jones, *Computer Programs for Infrared Spectrophotometry. Fourier Transform Programs* [#LI-LV], National Research Council of Canada Bulletin No. 18 (National Research Council, Ottawa, 1986).
14. J. K. Kauppinen, D. J. Moffatt, H. H. Mantsch, and D. G. Cameron, Appl. Spectrosc. **35**, 271 (1981).
15. H. H. Mantsch, H. L. Casal, and R. N. Jones, "Resolution Enhancement of Infrared Spectra of Biological Systems," in *Spectroscopy of Biological Systems*, R. J. H. Clark and R. E. Hester, Eds. (Wiley, New York, 1986), p. 1.
16. M. Levitt and J. Greer, J. Mol. Biol. **114**, 181 (1977).
17. M. Z. Papiz, L. Sawyer, E. E. Eliopoulus, A. C. T. North, J. B. C. Findlay, R. Sivaprasadarao, T. A. Jones, M. E. Newcomer, and P. J. Krauli, Nature **324**, 383 (1986).
18. L. Simons, G. Bergström, G. Blomfelt, S. Forss, H. Stenbäck, and G. Wansér, Comm. Physico-Math. **42**, 125 (1972).
19. H. Susi and D. M. Byler, Arch. Biochem. Biophys. **258**, 465 (1987).
20. L. K. Creamer, T. Richardson, and D. A. D. Parry, Arch. Biochem. Biophys. **211**, 689 (1981).
21. D. M. Byler and H. Susi, unpublished FT-IR data.
22. D. M. Byler, H. Susi, and H. M. Farrell, Jr., paper accepted by J. Dairy Sci. **71**, (1988).